

La vision : un modèle d'étude de la cognition

Mehdi SENOUSSE¹ & Laura DUGUÉ²

RÉSUMÉ. Nos différents sens – la vue, l'audition, le toucher, le goût, l'odorat – reçoivent constamment un flux massif d'informations. Toutes ces informations sont traitées et utilisées afin de guider nos actions. Les sciences cognitives représentent l'étude de ces facultés mentales par le prisme de différentes disciplines, e.g. linguistique, neuropsychologie, neuroscience ou modélisation. Chacune de ces disciplines considère les phénomènes mentaux et leur substrat physique, le système nerveux, comme un outil de traitement de l'information ayant pour but de guider le comportement de façon adaptative (Collins, Andler, & Tallon-Baudry, 2018). Les fonctions cognitives constituent ainsi une collection de systèmes de traitement de l'information servant différents buts, et dont les interactions sont à l'origine de la complexité de la cognition. L'étude de la cognition passe souvent par l'opérationnalisation de chacune de ces fonctions séparément. Par exemple, la mémoire permet de stocker et de réutiliser l'information, et l'attention permet de sélectionner celle qui est pertinente pour la tâche à effectuer, et d'en faciliter son traitement. Afin de caractériser les processus propres à une fonction cognitive donnée, il est alors nécessaire de fournir au sujet d'étude – ici nous nous concentrerons sur le primate humain et non-humain – une information à traiter, via différentes modalités sensorielles. Dans cet article d'opinion, nous nous concentrons sur la vision comme modèle d'étude singulier de la cognition à travers différents champs des sciences cognitives, de la psychologie cognitive aux neurosciences, en passant brièvement par la modélisation et la neuropsychologie. Notre objectif n'est pas de faire une description exhaustive de la modalité visuelle ni de faire une comparaison détaillée avec les autres modalités sensorielles, mais d'argumenter que l'accumulation des connaissances que nous en avons, ainsi que son organisation caractéristique du point de vue perceptif, algorithmique et physiologique, en font un modèle particulièrement riche de l'étude des fonctions cognitives. Après une brève présentation de certaines bases de la vision, nous illustrerons notre argument en nous concentrant sur une fonction cognitive spécifique : l'attention, et en particulier, son étude en psychologie cognitive et neurosciences. Nous aborderons notamment la façon grâce à laquelle nos connaissances sur la vision nous ont permis de comprendre les mécanismes comportementaux et neuronaux qui sous-tendent la sélection de l'information par l'attention, et la facilitation de son traitement. Nous concluons que les systèmes sensoriels peuvent être utilisés comme modèles d'étude de la cognition dans divers domaines des sciences cognitives.

Mots-clés : Attention, Cognition, Modèle, Modalités sensorielles, Vision

ABSTRACT. Vision: a model to study cognition.

Our senses – vision, audition, touch, taste and smell – constantly receive a large amount of information. This information is processed and used in order to guide our actions. Cognitive sciences consist in studying mental abilities through different disciplines, e.g. linguistic, neuropsychology, neuroscience or modelling. Each discipline considers mental phenomena and their physical substrate, the nervous system, as a tool to process information in order to guide behavior adaptively (Collins, Andler, & Tallon-Baudry, 2018). Cognitive functions are a collection of processing systems serving different goals, and whose interactions are key to the complexity of

¹ Department of Experimental Psychology, Faculty of Psychology and Educational Sciences, Ghent University, Ghent, Belgium. senoussi.m<at>gmail.com

² CNRS (Integrative Neuroscience and Cognition Center, UMR 8002) ; Université de Paris, Paris Descartes, et Institut Universitaire de France, Paris, France. laura.dugue<at>parisdescartes.fr

cognition. Studying cognition often implies operationalizing each of these functions separately. For example, memory allows to store and reuse information, and attention allows to select relevant information for the task at hand, and to facilitate its processing. To characterize the processes of specific cognitive functions, it is thus necessary to provide to the studied subject – here we concentrate on human and non-human primates – an information to be processed, through different sensory modalities. In this essay, we concentrate on vision as a unique model to study cognition through different fields of cognitive sciences, from cognitive psychology to neurosciences, mentioning also briefly modeling and neuropsychology. Our objective is not to do an exhaustive description of the visual system, nor to compare in detail vision with other sensory modalities, but to argue that the accumulation of evidence on the visual system, as well as its characteristic perceptual, algorithmic and physiological organization, make it a particularly rich model to study cognitive functions. After a brief presentation of some properties of vision, we will illustrate our argument focusing on a specific cognitive function: attention, and in particular its study in cognitive psychology and neuroscience. We will discuss how our knowledge of vision allowed us to understand the behavioral and neuronal mechanisms underlying attentional selection and facilitation of information. We will finally conclude that sensory systems can be used as models to study cognition in different fields of cognitive sciences.

Keywords: Attention, Cognition, Model, Sensory modalities, Vision

INTRODUCTION

Tous les jours, nous faisons l'expérience sensorielle de notre environnement. Nous visualisons le monde qui nous entoure ; nous touchons les objets qui le composent ; nous sentons les odeurs qui s'en dégagent ; nous entendons les sons qui sont émis ; nous goûtons les saveurs de ce que nous consommons. Il est donc nécessaire de traiter l'afflux constant d'informations sensorielles qui nous arrivent à chaque seconde. Toutes ces modalités sensorielles nous permettent d'interagir avec le monde et d'adapter notre comportement afin d'atteindre différents buts pertinents pour notre organisme. Par exemple, pour se rendre au travail, nous devons nous déplacer dans un environnement complexe, faire attention aux véhicules en traversant la route, se rappeler de passer à la boulangerie pour acheter à manger pour le déjeuner ou encore s'adapter aux potentiels problèmes sur le chemin, comme une panne des transports publics ou de son moyen de locomotion. Nous utilisons une gamme de fonctions mentales afin de réaliser toutes ces tâches comme l'attention, la mémoire ou la prise de décision. L'étude de la cognition s'intéresse à caractériser chacune de ces fonctions et de décrire les processus en jeu dans leur fonctionnement. Pour cela, les chercheurs en sciences cognitives, et en particulier les expérimentalistes, s'appuient sur les informations provenant de différentes modalités sensorielles pour étudier la façon dont l'information est traitée et transformée. Pour étudier la mémoire, nous pouvons demander à un participant de mémoriser une liste de chiffres apparaissant sur un écran (i.e. vision) ; pour l'attention, de détecter un son provenant de la droite en ignorant ceux provenant de la gauche (i.e. audition). De par leur nature, différentes modalités sensorielles présenteront différents avantages et inconvénients selon la question expérimentale posée. Par exemple, l'intrication physiologique du système olfactif avec le lobe temporal médian – intervenant dans l'encodage, le stockage et le rappel – en fait un système particulier pour l'étude de la mémoire (Eichenbaum, 1998). L'audition peut, elle, être un modèle d'étude intéressant de l'étude de l'attention. Des chercheurs ont notamment étudié l'effet *cocktail party* (Handel, 1989) dans lequel l'attention permet d'extraire l'information pertinente de la complexe superposition de bruits environnants. Cet effet est, par exemple, mis en évidence lorsque l'on porte son attention sur la parole d'un locuteur dans le

brouhaha ambiant d'une fête (d'où le terme « *cocktail party* »). Toutefois, si nous nous intéressons à l'étude de propriétés fines de l'attention spatiale, la vision est privilégiée de par son excellente résolution spatiale. Différentes modalités peuvent donc être choisies comme *modèle d'étude* d'une fonction cognitive. Dans cet article d'opinion, nous argumentons que la vision joue un rôle important en sciences cognitives car elle fut et est encore souvent choisie comme modèle d'étude principal. Nous n'aspérons pas à comparer la vision aux autres modalités, mais à argumenter que les connaissances que nous avons de son fonctionnement en font un modèle riche pour l'étude de certaines fonctions cognitives.

La vision joue un rôle majeur dans notre capacité à interagir avec notre environnement (e.g. lorsque nous naviguons dans l'espace, utilisons un outil, etc.) et s'avère être une fonction exacerbée chez toutes les espèces de primates diurnes. Nos grandes connaissances de cette modalité proviennent de nombreuses études en neurosciences et en psychologie. Les travaux d'Ernst Weber (1795-1878) et Gustav Fechner (1801-1887) ont, par exemple, permis de formaliser notre capacité à discriminer différents stimuli visuels. Ils ont développé des lois mathématiques mettant en lien l'intensité absolue d'une stimulation, par exemple visuelle ou tactile, avec l'intensité perçue par le sujet, et ont ainsi montré une relation logarithmique, et non linéaire, entre les deux. On pense également aux travaux fondateurs de Santiago Ramon y Cajal sur l'organisation cytoarchitectonique de la rétine (1894), et, plus tard, à l'étude du réseau d'interconnexions entre les différentes aires visuelles corticales de Daniel Felleman et David Van Essen (1991). Nous savons maintenant qu'une grande proportion de notre cerveau est dédiée au traitement de l'information visuelle, des estimations allant jusqu'à la moitié du cortex chez certains primates (Serenio, McDonald, & Allman, 1994). L'étude fonctionnelle du système visuel a aussi connu une avancée majeure avec les travaux de David Hubel et Torsten Wiesel (1959, 1962). Ces deux neurobiologistes, nobélisés en 1981, ont démontré que certains neurones sont sensibles à des traits visuels spécifiques orientés – il s'agit de la notion de *sélectivité à l'orientation* – et ont décrit comment cette sélectivité se construit à travers les étapes du traitement visuel. Ainsi, la sélectivité à des formes géométriques (e.g. un carré), par exemple, émerge de la minutieuse combinaison des projections de plusieurs types de neurones (e.g. sélectifs à des lignes horizontales et verticales) soulignant le lien critique en structure des connexions neuronales et fonction.

Au-delà de l'intérêt fondamental de comprendre les mécanismes de traitement de l'information visuelle, l'étude de la vision a permis de découvrir certaines propriétés universelles de la cognition, et plus spécifiquement, du système neuronal. Cela permet notamment de définir la notion de champ récepteur, qui peut être formalisé par la portion de l'espace sensoriel qui provoque une réponse pour un neurone donné. Ce concept fût défini initialement par Charles Sherrington (1906) comme étant une zone de la peau dont la stimulation peut provoquer un réflexe moteur dirigé vers cette même zone, e.g. un réflexe de grattage. Ce terme fut ensuite réutilisé par Haldan Hartline (1938) pour décrire la région du champ visuel induisant une réponse au niveau d'un neurone ganglionnaire spécifique de la rétine (voir section Vision pour plus de détails sur l'organisation de la rétine). Par la suite, ce terme a été étendu à d'autres neurones du système visuel, puis à d'autres systèmes sensoriels (e.g. audition, somatosensation ; Alonso & Chen, 2009). Progressivement, l'accumulation de connaissances sur la structure fonctionnelle du système visuel nous a permis de mieux comprendre les règles régissant la perception (Pelli, 2008), ainsi que des comportements complexes associés à la vision, tels que la lecture (Dehaene, Cohen, Morais, & Kolinsky, 2015) ou nos capacités d'imagination, aussi appelées imagerie mentale (Pearson, 2019).

Les découvertes sur l'architecture et les propriétés fonctionnelles des neurones du système visuel ont été d'une grande importance dans le développement de la modélisation cognitive et de

l'intelligence artificielle. Un exemple fondateur en modélisation cognitive, et plus précisément, dans l'étude des réseaux de neurones artificiels est le *perceptron* de Frank Rosenblatt publié en 1962. L'idée était de reproduire de façon algorithmique, puis dans une machine physique (un robot ou ordinateur), les capacités de catégorisation et de reconnaissance visuelle humaine, de caractères alphanumériques par exemple (Bishop, 2006). Quelques centaines de capteurs d'une caméra vidéo, représentant la rétine, étaient branchés à un réseau de neurones artificiels. Grâce à nos connaissances du système visuel et de sa plasticité, ainsi que les recherches en psychologie sur la reconnaissance visuelle et la lecture, l'algorithme de Frank Rosenblatt permit de construire un système permettant de discriminer des caractères alphanumériques (1962). Bien que limité à résoudre des problèmes simples dans sa version initiale, cette avancée dans la modélisation du comportement humain, et plus précisément de ses capacités visuelles, a été une des pierres fondatrices de la modélisation en sciences cognitives et de l'intelligence artificielle.

Au cours de la décennie suivante, les travaux de modélisation de David Marr s'inspirant, entre autres, des recherches sur la vision, eurent un impact important sur l'étude de la cognition. Il commença par proposer des théories computationnelles du fonctionnement du cerveau en s'intéressant à sa physiologie et son anatomie. Par théorie computationnelle, nous entendons une description des algorithmes de traitement de l'information obtenue grâce à leur implémentation physique, e.g. des neurones ou des transistors. Il se réorienta par la suite vers l'étude du système visuel et travailla à l'élaboration de modèles mathématiques de la perception visuelle. Ses travaux s'inspiraient tant des découvertes et progrès en psychologie expérimentale, informatique et intelligence artificielle, qu'en neuropsychologie et neurophysiologie (Stevens, 2012). Son approche multidisciplinaire permit à Marr de développer des théories computationnelles unificatrices qui furent particulièrement influentes en sciences cognitives, notamment concernant la perception des contours ou la vision en 3 dimensions. Plus encore, il énonça des paradigmes d'études généraux pour les sciences cognitives, i.e. les différents niveaux d'étude d'un système. Dans son paradigme, au niveau *computationnel*, on se demande quelle est la fonction remplie par le système ; au niveau *algorithmique*, quelles sont les représentations et les transformations permettant de remplir cette fonction ; et au niveau *d'implémentation*, comment ces algorithmes sont-ils implémentés au niveau physique (Marr, 2010). Bien que cette proposition de trois niveaux d'études soit encore discutée, elle demeure une étape majeure de l'histoire des sciences cognitives ; cela en partie grâce à la convergence de différentes disciplines (psychologie, intelligence artificielle, neurosciences, etc.), mais également grâce à un modèle d'étude commun : la vision. Les décennies suivantes virent fleurir de nombreux programmes de recherche en neurosciences computationnelles et intelligence artificielle, notamment dans le domaine du traitement automatique des langues, mais surtout en vision artificielle (LeCun, Bengio, & Hinton, 2015). Cette quête de formalisation des capacités de reconnaissance visuelle humaines a permis de nombreux progrès dans l'étude de la cognition par sa modélisation. Par exemple, un modèle de reconnaissance d'objet, nommé HMAX (Hierarchical Model And X), s'est inspiré de la structure hiérarchique entre certains neurones de l'aire visuelle primaire (V1) (Riesenhuber & Poggio, 1997 ; Serre, Oliva, & Poggio, 2007 ; Hubel & Wiesel, 1962). De nombreux autres modèles se sont inspirés de différentes propriétés neurophysiologiques du système visuel, telles que la forme des champs récepteurs ou la normalisation de l'activité d'un neurone par les neurones environnants (Carandini & Heeger, 2012 ; Heeger, Simoncelli, & Movshon, 1996 ; Marčelja, 1980). Ces modèles, qui purent être validés en les comparant aux performances comportementales et à l'activité cérébrale chez l'humain (Eickenberg, Senoussi, Pedregosa, Gramfort, & Thirion, 2013 ; Freeman & Simoncelli, 2011 ; Kay, Naselaris, Prenger, & Gallant, 2008 ; Thirion et al., 2006), ont

permis de mieux décrire le fonctionnement de la vision mais aussi de mieux comprendre des mécanismes généraux de la cognition, notamment d'apprentissage (Monier, 2018). Le biomimétisme, s'inspirer de la structure du vivant comme processus d'innovation (Bar-Cohen, 2006), tel que décrit plus haut, a ainsi été critique au développement de théories computationnelles de la cognition et aux avancées en intelligence artificielle.

Une des plus importantes avancées récentes en intelligence artificielle fut les réseaux de neurones profonds (en anglais *Deep Neural Networks*, DNN) qui consistent en une succession de couches d'unités représentant des neurones artificiels. L'avantage des DNN est qu'ils peuvent apprendre à accomplir de nombreuses tâches complexes sans devoir être finement paramétrés par l'utilisateur. Les paramètres de ces réseaux de neurones sont ajustés selon des lois mathématiques simples lors d'une phase d'apprentissage, ce qui leur permet ensuite d'accomplir les mêmes tâches sur de nouveaux stimuli. Les DNN sont ainsi capables de généraliser leur apprentissage (LeCun et al., 2015 ; LeCun, Bottou, Bengio, & Haffner, 1998). Développés notamment pour effectuer la reconnaissance d'objets visuels, ces modèles ont récemment été utilisés très largement car ils atteignent d'excellentes performances, souvent identiques et parfois supérieures à celles de l'être humain (e.g. détecter et catégoriser des objets : Krizhevsky, Sutskever, & Hinton, 2012 ; Szegedy et al., 2015). La structure de ces modèles s'inspire du système visuel biologique : des couches successives d'unités de traitement sensibles à des caractéristiques visuelles de plus en plus complexes. Des études récentes se sont intéressées aux similarités fonctionnelles entre le système visuel biologique et ses homologues artificiels. Elles ont montré une correspondance entre les représentations visuelles issues de ces modèles, apprises sans supervision humaine directe, et celles du cortex inféro-temporal, important dans la reconnaissance d'objets, chez l'humain et le primate non-humain (Eickenberg, Gramfort, Varoquaux, & Thirion, 2017 ; Güçlü & Gerven, 2015 ; Long, Yu, & Konkle, 2018 ; Yamins et al., 2014). Khaligh-Razavi et Kriegeskorte (2014) ont de plus montré que les DNN prédisant le mieux les réponses neuronales étaient ceux entraînés à effectuer des tâches pertinentes pour l'humain, e.g. distinguer des visages parmi d'autres objets. La structure du système visuel semblerait donc contrainte par une certaine fonction d'utilité comportementale, e.g. la capacité fondamentale de distinguer des visages d'êtres humains parmi des visages d'autres espèces animales est permise par des aires visuelles dédiées à une telle reconnaissance. Ces recherches montrent l'apport que nos connaissances du système visuel ont eu en modélisation et intelligence artificielle.

L'étude de la structure du système visuel a aussi permis de mettre en évidence des différences inattendues dans certaines pathologies. Des recherches récentes montrent par exemple que le comportement visuel (mouvements oculaires), les performances dans différentes tâches visuelles (e.g. recherche visuelle) et l'organisation neuronale du système visuel sont différents chez des personnes présentant des troubles du spectre autistique comparé à des personnes contrôles (Park et al., 2017 ; Robertson & Baron-Cohen, 2017). Il a par exemple été montré que le traitement du bruit présent dans des stimuli visuels était moins efficace chez les personnes présentant des troubles du spectre autistique (Park et al., 2017). Ces recherches ont de toute évidence été permises grâce à la somme de connaissances acquises sur la structure et le développement du système visuel, montrant comment l'utilisation de la vision comme modèle d'étude pourrait aider à mieux comprendre et étudier les troubles du spectre autistique. Il s'agit là encore d'un autre exemple pour lequel les sciences cognitives, ici via l'approche neuropsychologique, ont bénéficié de l'utilisation de la modalité visuelle comme modèle d'étude. La vision peut ainsi jouer le rôle de modèle d'étude de la cognition pathologique. Grâce à la grande quantité de connaissances sur cette modalité

sensorielle, il est possible de mieux comprendre la nature des dysfonctionnements présents dans différentes pathologies.

Ces découvertes sur le fonctionnement de la vision nous ont permis d'en faire de nombreuses autres. Dans cet article, nous argumentons que, grâce à ses propriétés particulières, il est possible de tirer des conclusions au-delà de cette modalité sensorielle, et plus généralement à l'étude de la cognition en l'utilisant comme modèle d'étude pour mieux comprendre certaines fonctions cognitives ou certaines pathologies.

LA VISION

L'étude de la vision a parcouru un long chemin depuis la théorie de l'émission (ou extramission) qui postulait qu'elle reposait sur l'émission, par l'œil, de rayons lumineux « saisissant » les objets pour les percevoir (Empedocles et Platon, 5s. avant J.C.). De nombreux philosophes, médecins et, plus tard, psychologues, neuroscientifiques et même physiciens se sont penchés sur le système visuel. Dans cette section, nous reprendrons quelques faits marquants de l'histoire de l'étude de la vision. Puis, nous présenterons certains principes d'organisation du système visuel qui nous permettront de mieux comprendre comment celui-ci peut servir de modèle d'étude de fonctions cognitives au-delà de la perception.

Quelques bases et un peu d'histoire

Les premières mentions du cerveau remontent à un papyrus datant de 1700 avant J.C., qui serait lui-même la copie d'un manuscrit datant de 3000 avant J.C. (Breasted, 1930 ; Gross, 1999). Plusieurs cas de lésions neurologiques sont décrits et constituent les plus anciens rapprochements connus entre cerveau et pensée. Il fallut attendre Alcmeon de Croton (environ 500 avant J.C.), médecin et philosophe, pour voir naître la thèse que le cerveau était le siège de la sensation et la cognition (pour un historique détaillé, voir Gross, 1999). D'intérêt particulier pour notre article, lors de ses dissections, Alcmeon de Croton s'était intéressé tout particulièrement aux sens, et plus précisément à la vision. Il décrivit les nerfs optiques et avança l'idée qu'ils étaient des chemins permettant de transporter la lumière au cerveau (Gross, 1999). Une des grandes questions qui a mêlé physiologie, mathématiques, physique et psychologie fut l'étude de l'optique géométrique pour comprendre les phénomènes de réflexion et réfraction de la lumière. Après les prémisses des lois de l'optique, entre autres par Euclide, le mathématicien perse Alhazen dément la théorie de l'émission, où l'œil serait source de lumière qui permettrait de percevoir. Il en conclut : « [...] l'œil ne perçoit pas comme visible tout objet situé dans la même ambiance, cette perception ne peut se réaliser uniquement que par la réflexion de la lumière et en suivant des lignes droites qu'on peut prolonger par l'imagination entre la surface de l'objet et la surface de l'œil. ». Ces travaux sur la lumière et la vision, puis de nombreux autres en optique (René Descartes, Isaac Newton, etc.), ont permis d'établir les fondements de la psychologie expérimentale de la vision.

Au 19^{ème} siècle, Ernst Weber et Gustav Fechner établirent les bases de la psychophysique, i.e. l'étude de la relation quantitative entre la quantité physique d'un stimulus (e.g. luminance, taille, poids) et la sensation ou la perception qu'il produit. Gustav Fechner publie le livre "*Elemente der Psychophysik*" en 1860, qui sera un ouvrage fondateur de la psychologie. Il y décrit des théories et techniques classiques de la psychophysique, parmi lesquelles les méthodes d'estimation des seuils sensoriels de détection et de discrimination, ou encore l'évolution logarithmique de l'intensité perçue d'un stimulus en fonction de son intensité physique. Ces travaux, ainsi que ceux de Hermann von Helmholtz, Wilhelm Wundt ou encore, plus tard, William James, permirent de fonder l'étude scientifique de la psychologie.

Concernant les substrats physiques du système visuel, une découverte majeure, bien que ce ne fut compris que plus tard, fut la *strie de Gennari*, au 18^{ème} siècle, par Francisco Gennari (Gross, 1999). Cette ligne blanche le long de l'épaisseur du cortex représente une épaisse bande d'axones myélinisés dans une des couches corticales de V1, et permet de localiser cette aire visuelle qui de nos jours est probablement la région la mieux décrite du cortex. Quelques décennies plus tard, différentes personnes firent la découverte de la propriété rétinotopique de V1 : le fait que l'aire visuelle primaire contienne une carte entière du champ visuel, conservant ainsi la topographie de ce dernier (Henschen, 1893 ; Holmes & Lister, 1916 ; Inouye, 1909). L'étude de V1 fut, au milieu du 20^{ème} siècle, une des plus grandes avancées scientifiques en biologie et neurosciences grâce aux travaux de David Hubel et Torsten Wiesel (1959, 1962), qui leur a valu le prix Nobel en 1981. Leurs travaux sont décrits plus amplement dans la section « De simple à complexe : une hiérarchie fonctionnelle ».

Les études de cas utilisant la méthode anatomo-clinique, prémisses de la neuropsychologie, ont permis de comprendre une autre échelle de l'organisation du système visuel. L'un des premiers cas de déficit visuo-moteur a été signalé par Rezso Bálint en 1909 chez un patient ayant subi des lésions des lobes pariétaux postérieurs. Ce patient était incapable d'atteindre avec précision des objets distaux mais pouvait indiquer précisément différentes parties de son corps les yeux fermés, excluant une déficience strictement motrice (Rossetti, Pisella, & Vighetto, 2003). Ces études de cas de Bálint, suivies de celles de Ungerleider et Mishkin (1982) chez le primate non-humain, et de Milner et Goodale (1994) chez des patients humains ont permis d'attribuer au cortex occipito-pariétal des fonctions spatiales et visuo-motrices. Des déficits dans la perception de la couleur ou de la forme des objets, ont, quant à elles, été observées lorsque le cortex occipito-temporal était lésé. Mishkin et Ungerleider (1983) ont montré une perte de la capacité à discriminer des objets après une ablation d'une zone du cortex temporal inférieur. Certains patients humains souffrant de lésions dans cette même zone montraient également ce type de déficits (Feinberg & Farah, 2000). L'idée de cette division neuronale des fonctions visuelles a été formulée dès les années 70 par Gerald Schneider (1969) et fut mise à jour par Ungerleider et Mishkin qui proposèrent les voies désormais célèbres du « quoi », pour la reconnaissance, et du « où », pour la localisation dans l'espace. Milner et Goodale ont ensuite proposé que la voie du « où » serait une voie visuo-motrice servant à interagir avec des objets. L'hypothèse des deux voies visuelles, plus récemment agrémentée d'une troisième voie du « quand » (Battelli, Pascual-Leone, Cavanagh, 2007) a été très influente et reste un cadre théorique important pour comprendre comment la vision est organisée et comment différentes fonctions perceptives interagissent.

Toutes ces découvertes ont fait du système visuel l'un des mieux décrit en comparaison des autres modalités sensorielles. De plus, la structure rétinotopique, les interconnexions entre ses différentes aires et sa structure fonctionnelle (voies dorsales et ventrales) en font un modèle d'étude particulièrement riche pour étudier les fonctions cognitives telles que la mémoire, l'apprentissage, l'attention ou la décision, et comprendre comment ces fonctions se servent, sont contraintes et/ou modifient cette organisation. Voyons maintenant plus en détails certaines propriétés du système visuel avant de prendre l'exemple de l'étude de l'attention pour argumenter que les propriétés de la vision en font un modèle d'étude particulièrement riche de la cognition.

De la lumière à l'activité cérébrale

Pour comprendre comment le cerveau traite l'information visuelle, passons en revue les premières étapes de transformation de la lumière depuis la rétine jusqu'au cerveau.

Les informations visuelles sont captées par nos rétines, ensembles de couches de cellules tapissant le fond de chaque œil et composées notamment de photorécepteurs (les cônes et les bâtonnets). Ces cellules sont sensibles à la lumière pénétrant nos yeux après avoir été réfléchiée par des objets environnants. Elles réagissent à ce rayonnement électromagnétique en modifiant leur potentiel de membrane. Ce potentiel est transmis aux cellules ganglionnaires qui le relaient au cerveau par leurs axones. C'est cet ensemble de fibres transportant l'information recueillie par la rétine qui constitue le nerf optique. De façon intéressante la partie de la rétine d'où partent ces fibres vers le cerveau ne contient pas de photorécepteurs et est donc « aveugle », elle est appelée « la tâche aveugle » ou *blind-spot* en anglais. Malgré l'absence d'information rétinienne dans cette partie du champ visuel nous ne voyons pas de « trou » dans l'image consciente de la scène visuelle. Ce phénomène a été l'objet de nombreuses études afin de comprendre comment la perception visuelle consciente se construit malgré cette tâche aveugle (He & Davis, 2001). Les fibres constituant le nerf optique se croisent au niveau du chiasma optique afin de véhiculer les informations provenant de chaque hémichamp visuel (droit et gauche), collectés par les deux yeux, à l'hémisphère cérébral opposé. Ces informations rejoignent (principalement) le corps genouillé latéral (CGL) dans le thalamus. Depuis cette première étape de traitement dans le CGL, des projections sont finalement envoyées au cortex visuel.

Organisation rétinotopique

L'information visuelle est structurée selon plusieurs principes, parmi lesquels le principe fondamental de la *rétinotopie*. La rétinotopie désigne le fait que certaines aires cérébrales représentent les informations visuelles en respectant l'arrangement spatial des capteurs de la rétine. Cette représentation conserve donc, à un coefficient de transformation près, les distances et la structure spatiale de la scène visuelle. Ainsi, deux points proches dans le champ visuel seront représentés par deux populations de neurones proches dans une aire rétinotopique. Cette organisation est tout d'abord observée dans V1, puis dans de nombreuses aires de la hiérarchie visuelle (Felleman & Van Essen, 1991 ; Nowak & Bullier, 1997). Cette organisation cartographiée de l'espace visuel est également caractérisée par des divisions de l'espace visuel sous forme de quadrant. Ainsi, chaque hémisphère cérébral représente l'information visuelle du champ opposé grâce à la décussation du nerf optique, i.e. l'aire V1 de l'hémisphère gauche représente l'hémichamp visuel droit. De la même manière, la partie dorsale des aires visuelles V1, V2 et V3 représentent le quadrant inférieur de l'hémichamp opposé alors que la partie ventrale représente le quadrant supérieur (Wandell & Winawer, 2011). L'arrangement rétinotopique des aires visuelles a en retour des conséquences sur la perception. Il a par exemple été montré que l'effet de *crowding*, la dégradation des performances de reconnaissance visuelle d'un objet lorsque d'autres sont situés à proximité dans le champ visuel, serait déterminée par la proximité spatiale de ces objets dans V1 (Pelli, 2008). Certaines lois régissant notre perception, et certains comportements, découlent donc de la structure physiologique de V1.

Tout au long de cet acheminement des informations visuelles, les neurones partagent un mécanisme fonctionnel commun : une sensibilité à des caractéristiques visuelles spécifiques dans une certaine partie du champ visuel, appelée *champ récepteur*. Dans toutes les aires visuelles rétinotopiques, la taille de ces champs récepteurs augmente avec l'excentricité. Ce principe est nommée *magnification corticale* (Baizer, Ungerleider, & Desimone, 1991). De plus, la taille des champs récepteurs augmente, pour une même excentricité, à travers la hiérarchie visuelle, de V1 à V2, puis de V2 à V3, etc. Malgré la grande similitude de structure et fonction à travers les primates d'une même espèce, il existe tout de même une certaine variabilité inter-individuelle,

notamment dans la taille de ces aires rétinitopiques. Il a par exemple été montré que la taille de V1 influence des caractéristiques objectives et subjectives de l'imagerie mentale (Bergmann, Genç, Kohler, Singer, & Pearson, 2016).

De simple à complexe : une hiérarchie fonctionnelle

Les neurones des aires rétinitopiques du système visuel ont une sélectivité à certains traits visuels au sein de leurs champs récepteurs. La caractéristique la plus étudiée des neurones de V1 est ainsi la sélectivité à des barres orientées contrastées, alors que les neurones de V4 sont sélectifs, entre autre, à des conjonctions de contours formant des angles. Cette sélectivité devient de plus en plus complexe à travers la hiérarchie des aires visuelles. La notion de traitement hiérarchique dans le cortex visuel a été avancée par Hubel et Wiesel alors qu'ils étudiaient V1 (Hubel & Wiesel, 1962). Tout en décrivant les propriétés physiologiques des cellules de V1 à travers les six couches composant le cortex, ils mirent en évidence une augmentation de la complexité de leurs champs récepteurs et de leurs sélectivités. Depuis la couche 4 du cortex visuel primaire recevant les projections du CGL, jusqu'aux couches 2/3 d'où partent les projections vers les aires rétinitopiques adjacentes (e.g. V2, V3), une invariance se développe quant aux types de caractéristiques visuelles auxquelles les neurones sont sensibles. Tout d'abord, V1 reçoit des projections du CGL en deux systèmes principaux distincts : parvocellulaire (au niveau des couches 3 à 6), qui assurerait la vision des détails et des couleurs, et magnocellulaire (couches 1 et 2), qui assurerait la vision du mouvement. Puis, pour décrire l'évolution fonctionnelle à travers la hiérarchie visuelle, Hubel et Wiesel ont initialement montré que certains neurones de V1 répondaient à la présentation de barres lumineuses ayant une certaine orientation, chaque neurone ayant son orientation préférée – il s'agit des cellules simples (Hubel & Wiesel, 1959). Ils montrèrent également que la sélectivité des cellules simples émerge de l'agrégation des projections de plusieurs neurones du CGL ; et, de la même manière, l'information provenant de plusieurs cellules simples converge vers des cellules traitant des informations plus complexes – il s'agit des cellules complexes. La sélectivité de ces cellules aux contrastes devient tolérante à des changements de position. En d'autres termes, alors que les cellules simples sont sélectives à des contrastes à un endroit précis de leur champ récepteur, les cellules complexes, qui agrègent la réponse de plusieurs cellules simples, sont sélectives à ces contrastes quelle que soit leur position au sein d'un champ récepteur plus large. Cette progression dans la complexité de la sélectivité des neurones, du CGL aux cellules simples, puis complexes de V1 est un premier élément contribuant à la structure hiérarchique du système visuel.

L'organisation hiérarchique du traitement visuel, et son complexe réseau de connexions entre les différentes aires, ont par la suite été montrés à l'aide de différentes techniques telles que la cytoarchitectonie ou le traçage (Felleman & Van Essen, 1991). Il a été montré plus tard, grâce à des modèles computationnels, que cette architecture *en avant* (en anglais *feed-forward*) peut accomplir des tâches complexes de reconnaissance visuelle (Serre, Oliva, & Poggio, 2007 ; VanRullen, Delorme, & Thorpe, 2001). Sa formulation stricte en tant que processus en avant est néanmoins incomplète car il existe de nombreuses boucles de rétroaction à différents niveaux de la hiérarchie (e.g. neurones de V4 projetant aux neurones de V1).

L'étude des propriétés fonctionnelles de la hiérarchie des aires visuelles, c'est-à-dire des traits auxquels les neurones des aires au-delà de V1 sont sensibles, s'avère être plus difficile. En utilisant une méthode d'analyse permettant d'estimer la distribution de la sélectivité des neurones de V4, David et al. (2006) ont montré que, comparés à V1, les neurones de V4 étaient sélectifs à des stimuli plus complexes en termes d'orientation et de fréquences spatiales. Par exemple, les

neurones de V4, au contraire de ceux de V1, ont une sélectivité de fréquence spatiale bimodale, i.e. sélectifs à plusieurs orientations. Certains neurones de V4 sont, de part cette double sélectivité, sélectifs à la conjonction de lignes formant différents angles et permettant de traiter les contours d'un objet. Pasupathy et Connors (2002) ont par ailleurs montré que des populations de neurones dans l'aire V4 encodent les caractéristiques « principales » des contours d'objets visuels, e.g. concave vers la droite. Ces résultats montrent que V4 permet de représenter des objets complexes par une somme de leurs parties constituantes.

La complexité des caractéristiques auxquelles les neurones sont sélectifs augmente le long de la hiérarchie des aires visuelles. Il devient alors de plus en plus difficile de formaliser leur sélectivité. Pour comprendre l'organisation du système visuel, il a aussi été utile de prendre le problème dans l'autre sens : au lieu de caractériser la sélectivité des neurones de différentes aires visuelles à l'aide de stimuli de plus en plus complexes, des chercheurs ont étudié comment le système visuel répond à des stimuli naturels (visages, voitures, maisons, etc.). Ainsi, il a été possible, dès les années 70, de montrer chez le primate non-humain une sélectivité des neurones dans le cortex inféro-temporal (IT) aux visages et à d'autres objets (Gross, Rocha-Miranda, & Bender, 1972). Par la suite, en utilisant des techniques de neuroimagerie chez l'humain, il a été mis en évidence l'existence de différentes régions d'IT répondant à des groupes de stimuli appartenant à une même catégorie d'objets (e.g. objets, lettres, visages). Spécifiquement, plusieurs zones de sélectivité ont été établies à l'aide de l'imagerie par résonance magnétique fonctionnelle (IRMf). Par exemple, une zone sélective aux visages dans le gyrus fusiforme droit (Face Fusiform Area, FFA ; Kanwisher, McDermott, & Chun, 1997), une autre sélective aux lieux et aux scènes dans le cortex parahippocampique (Parahippocampal Place Area, PPA ; Epstein & Kanwisher, 1998), ou encore aux caractères écrits dans le gyrus fusiforme gauche (Visual Word Form Area, VWFA ; Cohen et al., 2000) et aux parties du corps dans le cortex occipital latéral (Extrastriate Body Area, EBA ; Downing, Jiang, Shuman, & Kanwisher, 2001). Cette dichotomie entre aires rétinitopiques et aires fonctionnelles tient plus à la chronologie de ces découvertes qu'à une réelle distinction fondamentale. En effet, les régions fonctionnelles, sélectives par exemple aux visages, conservent des propriétés rétinitopiques bien que moins précises, que les régions strictement rétinitopiques (e.g. V1, V2, V3 ; Kay, Weiner, & Grill-Spector, 2015 ; Rajimehr, Bilenko, Vanduffel, & Tootell, 2014 ; Wandell & Winawer, 2011).

Comme nous avons pu le voir dans ce bref aperçu, le système visuel a été étudié sous une multitude de facettes grâce à la psychologie, la neurologie ou encore la physiologie. Toutes ces recherches sur le système visuel n'avaient pas toujours pour but de mieux comprendre la vision en soi, mais pouvaient également s'intéresser à la formalisation mathématique des processus mentaux, en étudiant la transformation de l'intensité d'un stimulus en sensation chez Weber et Fechner, ou comprendre l'unité fondamentale du système nerveux (le neurone) lorsque Santiago Ramon y Cajal étudia la rétine. L'étude des fonctions cognitives comme la mémoire ou l'attention a, elle aussi, très tôt utilisé la vision comme modèle d'étude. Parmi les très célèbres expérimentalistes, Hermann von Helmholtz utilisait dès la fin du 19^{ème} siècle, un appareil expérimental (une boîte en bois) lui permettant d'étudier l'orientation de l'attention dans l'espace visuel en séparant celle-ci des mouvements oculaires (Carrasco, 2011). Le choix d'Helmholtz était probablement guidé par la simplicité technique et pratique de la manipulation. Quoi qu'il en soit, les connaissances acquises sur la vision ont depuis été d'une importance majeure dans la compréhension de l'attention. C'est ce que nous allons voir dans la partie suivante.

L'ATTENTION SÉLECTIONNE

L'attention : quelques définitions

Dès que nous ouvrons les yeux, nous sommes inondés par une grande quantité d'informations sensorielles. Chacune de nos actions, que ce soit lire un livre, faire la cuisine, ou traverser la rue, nécessite que nous analysions les informations qui sont pertinentes pour la tâche à réaliser. L'*attention* est la fonction cognitive qui permet de *sélectionner* l'information, pour en *faciliter* son traitement. Elle permet d'améliorer la perception et de guider notre comportement. Mais pourquoi ne pouvons-nous pas traiter toutes les informations qui sont à notre disposition ? Quel est le principe neurophysiologique rendant cette sélection nécessaire ? Il a été proposé que le traitement de l'information est contraint par une quantité d'énergie disponible pour le cerveau limité, et un coût énergétique de l'activité neuronale particulièrement élevé (Carrasco, 2011 ; Lennie, 2003). Si, à chaque moment dans le temps, la quantité d'énergie disponible ne varie pas, mais la demande énergétique, elle, est modulée, alors un système doit être mis en place pour allouer l'énergie en fonction de la demande. L'attention est donc sélective du fait de ressources énergétiques limitées.

L'attention, dont la terminologie peut parfois paraître complexe, ne doit pas être confondu avec « concentration » ou « vigilance » dont la dynamique temporelle est plus lente (i.e. plusieurs minutes, voire plusieurs heures). Elle donne la priorité à certaines informations au détriment d'autres dans le but d'améliorer notre traitement perceptif (Carrasco, 2011). Ce phénomène peut être mis en place en quelques centaines de millisecondes seulement. Nous pouvons porter attention à un endroit de l'espace (attention *spatiale*), à un objet (attention aux *objets*) ou à une caractéristique d'un objet (attention aux *traits*). Cette orientation attentionnelle peut se faire en dirigeant notre regard vers le stimulus en question (attention *ouverte*, i.e. l'attention est directement définie comme l'endroit où se pose, ouvertement, notre regard), ou bien en l'absence de mouvement des yeux ou de la tête (attention *couverte*). Nous pouvons diriger notre attention de manière volontaire (attention volontaire ou *endogène*). Notre attention peut à l'inverse être attirée par l'apparition soudaine d'un stimulus (attention involontaire ou *exogène*).

Toutes ces définitions, et il en existe d'autres, ont été établies grâce à de nombreuses années d'études qui ont connu un essor considérable à partir de la fin des années 80, et dont le modèle d'étude principal est le système visuel. L'objectif de cet article d'opinion n'est pas de faire une revue de la littérature sur l'attention – pour revue voir (Carrasco, 2011 ; Nobre & Kastner, 2014) dans lesquelles nous puisons d'ailleurs certains des exemples présentés plus bas – mais de démontrer par une série d'exemples que le système visuel est un modèle de choix de l'étude de l'attention.

Imager la vision pour mesurer l'attention

Si la quantité de ressources que l'on peut allouer à un objet ou un endroit de l'espace (que nous simplifierons dans la suite de l'article en parlant de *stimulus*) est constante, cela implique nécessairement que, lorsque notre attention est dirigée sur un stimulus, cela se fera au détriment d'autres stimuli. Des études ont en effet confirmé un bénéfice comportemental pour le stimulus attendu et un coût pour un stimulus inattendu. Comment nos connaissances du système visuel nous permettent-elles d'étudier les bases neuronales d'un tel compromis comportemental ? Dans la suite, nous allons donner une série d'exemples illustratifs répondant à cette question.

Les études fondatrices s'intéressant aux bases neuronales de fonctions cognitives, et en particulier de l'attention, ont utilisé l'électrophysiologie, méthode d'enregistrement invasive de l'activité neuronale, chez le primate non-humain (e.g. Moran & Desimone, 1985 ; Reynolds,

Pasternak, & Desimone, 2000). Grâce à cet outil, des auteurs ont établi les premières démonstrations du compromis entre coût et bénéfice créé par la sélection attentionnelle. Le principe était le suivant : des électrodes étaient implantées chez le macaque, dans des aires visuelles telles que V1, V2 et V4. Pour chaque site d'enregistrement, le champ récepteur d'un neurone ou groupe de quelques neurones (champ récepteur multi-unitaire), était cartographié, i.e. la région de l'espace visuel pour laquelle le groupe de neurones produit une activité. Pour chaque site d'enregistrement était ensuite déterminé le stimulus préféré, c'est-à-dire le stimulus produisant une activité neurale maximale permettant de caractériser la *sélectivité* des neurones enregistrés. Les stimuli testés correspondaient à une combinaison de traits simples (barres orientées de différentes couleurs) qui sont analysés par les aires visuelles précoces dans la hiérarchie visuelle (i.e. V1-V4). Les chercheurs ont montré que lorsque deux stimuli, un stimulus préféré et un autre non-préféré, étaient présentés dans le même champ récepteur, l'activité globale du groupe de neurones diminuait. Curieusement, lorsque le macaque portait attention au stimulus préféré, la réponse évoquée par la paire de stimuli augmentait (bénéfice), tandis que lorsque son attention était dirigée sur le stimulus non-préféré, la réponse évoquée diminuait (coût). La conclusion proposée par les auteurs est que la présence de plusieurs stimuli active des populations de neurones concurrentes et que l'attention sélectionne l'information en biaisant la compétition en faveur du stimulus attendu. En d'autres termes, l'activité des neurones sélectifs au stimulus attendu est favorisée, au détriment des autres neurones. L'attention est ainsi capable d'utiliser la structure du système visuel et d'en manipuler sa fonction pour faciliter le traitement perceptuel. Il est évident dans ce premier exemple que nos connaissances du système visuel (champs récepteurs et sélectivité) ont été essentielles à l'élaboration d'un protocole expérimental contrôlé, et ont permis une avancée majeure dans nos connaissances des bases neuronales de la sélection attentionnelle.

A la fin des années 80, l'IRMf, technique d'imagerie non-invasive et non-irradiante (contrairement à d'autres techniques telles que la TEP : tomographie par émission de positons), est venue augmenter la boîte à outils des chercheurs en neuroscience et psychologie cognitive. L'IRMf a ainsi permis d'imager le système visuel avec une résolution spatiale sans précédent. Au début des années 2000, les cartes rétinotopiques du cortex visuel du primate humain ont atteint une résolution spatiale d'environ deux millimètres (pour revue : Wandell & Winawer, 2011). En 1999, Brefczynski et DeYoe ont publié un article mettant à profit cette avancée technologique pour étudier les bases neuronales de la sélection d'une région de l'espace par l'attention couverte spatiale. Ils ont comparé les cartes rétinotopiques de plusieurs participants, obtenues grâce à la présentation de petits stimuli sur différentes régions de l'écran, aux cartes obtenues lorsqu'un large stimulus couvrant tout le champ visuel est présenté aux participants, qui sont instruits de déplacer leur attention d'un segment à l'autre de ce stimulus. Ils ont observé que les cartes rétinotopiques visuelles étaient identiques aux cartes purement basées sur l'attention spatiale. Ils ont ainsi pu créer les premières *cartes rétinotopiques de l'attention* (Brefczynski & DeYoe, 1999). Cette expérience suggère que l'attention est capable de sélectionner une région de l'environnement visuel en activant spécifiquement les régions visuelles traitant cette portion de l'espace. Cette démonstration a de plus permis d'établir une nouvelle méthode d'étude en IRMf des mécanismes neuraux sous-tendant la sélection attentionnelle : les analyses par régions visuelles d'intérêt (e.g. Dugué, Merriam, Heeger, & Carrasco, 2017 ; Liu, Pestilli, & Carrasco, 2005 ; Pestilli, Carrasco, Heeger, & Gardner, 2011). Il ne s'agit plus de rechercher quelle région est impliquée dans une tâche attentionnelle donnée. Dans ce cas, le chercheur enregistre l'activité du cerveau de nombreux participants au cours de la tâche, et moyenne l'activité à travers les participants pour révéler les régions significativement actives, sans hypothèse *a priori* quant à la région d'intérêt. Dans la

méthode par région d'intérêt, la question est maintenant de savoir comment la région est modulée par l'attention. Ici, la région ciblée est cartographiée pour chaque participant, et l'activité dans la région est analysée en fonction des conditions expérimentales. Cette approche permet une analyse fine de l'effet étudié au niveau individuel. Par ailleurs, elle permet de tenir compte des variabilités de taille et de position de la région corticale qui pourraient diminuer l'effet lorsque l'on moyenne l'activité cérébrale de tous les participants (Dugué et al., 2017).

La propriété de rétinotopie des aires visuelles a été mise à profit dans de nombreuses études, et notamment dans des expériences de stimulation magnétique transcrânienne (SMT) qui a la capacité de générer des phosphènes. Un phosphène est un flash illusoire produit par la stimulation des neurones du cortex visuel (V1 principalement). Il apparaît dans le champ visuel en coordonnée rétinotopique spécifique de la zone stimulée. C'est une manifestation physiologique d'une population de champs récepteurs. Dugué et VanRullen ont développé un protocole expérimental utilisant la cartographie des phosphènes dans le but d'obtenir un contrôle unique de l'effet de la stimulation (Dugué, Marque, & VanRullen, 2011, 2015 ; Dugué & VanRullen, 2017). Pour chaque participant, le phosphène correspondant à la zone stimulée était dessiné à l'écran. Cette zone était par la suite utilisée pour y présenter des stimuli visuels. Le contrôle expérimental correspondait à la présentation des stimuli dans la zone symétrique du phosphène dessiné par rapport au point de fixation, donc non affectée par la stimulation. En d'autres termes, pour une même stimulation, la SMT pouvait effectivement interagir avec le traitement des stimuli visuels, lorsqu'ils étaient présentés dans la zone du phosphène, ou non, lorsqu'ils étaient présentés dans la zone symétrique contrôle. Cette méthode a ensuite été utilisée pour étudier la dynamique temporelle de la sélection attentionnelle (Dugué, Roberts, & Carrasco, 2016). Dans une tâche classique d'indication attentionnelle (Posner, Snyder, & Davidson, 1980), les participants devaient rapporter l'orientation d'un stimulus attendu ou inattendu, alors qu'ils étaient stimulés par SMT à différents délais au cours de l'essai. Grâce à la cartographie des phosphènes de chaque participant, soit le stimulus attendu, soit le stimulus inattendu, était stimulé. Les auteurs ont ainsi mis en évidence que l'attention sélectionnait alternativement chacun des deux stimuli en faisant intervenir V1, et plus spécifiquement les populations de champs récepteurs de V1 codant pour les stimuli présentés, à la fréquence thêta ($\sim 5\text{Hz}$). Ce résultat a par la suite été répliqué dans une étude comportementale (Senoussi, Moreland, Busch, & Dugué, 2019). Cette série d'études a ainsi permis de démontrer que l'attention échantillonne l'information visuelle de façon périodique à la fois dans l'espace et dans le temps. Là encore, il s'agit de la propriété rétinotopique du système visuel, et le développement d'un protocole expérimental contrôlé, qui ont permis de mieux comprendre la dynamique temporelle de la sélection attentionnelle.

L'ATTENTION FACILITE

L'attention module les traitements visuels précoces

L'expérience visuelle débute par la décomposition de la scène en une somme de traits simples, tels que le contraste, l'orientation ou la couleur. Ainsi, un stimulus couramment utilisé en psychophysique est le *Gabor*. Le Gabor est un réseau sinusoïdal convolué avec une Gaussienne, présentant des barres orientées alternées noires et blanches ou colorées. L'avantage est que l'on peut changer ses caractéristiques, i.e. son contraste, son orientation ou sa taille, de manière paramétrique. Ce stimulus a été utilisé notamment dans des tâches comportementales de discrimination d'orientation, et a permis de démontrer que les performances dans ces tâches augmentent (e.g., Nachmias, 1967 ; Skottun, Bradley, Sclar, Ohzawa, & Freeman, 1987) en suivant une fonction sigmoïde (fonction de réponse au contraste), alors que la réponse IRMf augmente de

façon monotone (Boynton, Demb, Glover, & Heeger, 1999) avec l'augmentation du contraste du Gabor. Des chercheurs se sont ainsi posés la question de savoir si l'attention pouvait faciliter la perception en modulant la réponse comportementale au contraste (pour revue : Carrasco, 2011). Ils ont constaté que l'attention des participants dirigée sur le Gabor décalait la fonction de réponse au contraste soit vers la gauche (diminution du contraste nécessaire pour produire 50% de réponses correctes ; *gain de contraste*), soit vers le haut (augmentation des performances maximales à fort contraste ; *gain de réponse*). Ultérieurement, il a été rapporté que le ratio entre la taille du focus attentionnel et la taille du stimulus attendu est le facteur déterminant entre gain de contraste et de réponse (Herrmann, Montaser-Kouhsari, Carrasco, & Heeger, 2010 ; voir Reynolds & Heeger, 2009, pour modèle mécanistique). Cet exemple suggère que l'attention facilite la perception en modulant la sensibilité du système visuel au contraste. Cela pourrait laisser entendre que l'attention est capable de changer l'apparence subjective du contraste.

La question de savoir si l'attention est capable de changer l'apparence d'un stimulus n'est pas seulement un débat philosophique (Block, 2010). Elle date des pionniers de la psychologie expérimentale, tels que James, Fechner ou Helmholtz. L'équipe de Marisa Carrasco a plus tard développé le premier paradigme psychophysique rigoureux et objectif visant à étudier les effets de l'attention sur l'expérience subjective d'un stimulus visuel. Le Gabor et la sensibilité au contraste du système visuel ont été utilisés pour étudier les mécanismes sous-tendant ce possible changement subjectif (Carrasco, Ling, & Read, 2004). Deux Gabors étaient présentés aux participants alors que leur attention spatiale exogène (involontaire) était manipulée grâce à un indice visuel périphérique. Le contraste de chaque Gabor changeait indépendamment l'un de l'autre entre essais, et les participants devaient rapporter l'orientation du Gabor ayant le plus fort contraste. Les résultats indiquent que lorsque l'attention est attirée sur un stimulus donné grâce à un indice exogène, l'orientation du stimulus est mieux discriminée, et son contraste est subjectivement perçu comme plus haut qu'il ne l'est en réalité. Cette expérience démontre que l'attention altère la perception subjective du contraste. Par la suite, d'autres études ont démontré que l'attention altère également la perception de la taille d'un stimulus, de sa position, de son mouvement, de la distance entre stimuli, et d'autres traits visuels de bas niveau (pour revue Carrasco & Barbot, 2019). Il semblerait donc que l'attention facilite la perception en modulant l'activité neuronale des aires visuelles précoces (e.g. V1, V2, V3, etc).

L'attention module les aires visuelles précoces

Dans la section précédente, nous avons vu qu'il est possible d'isoler des paramètres visuels simples (e.g. contraste, orientation, etc) et de les moduler paramétriquement, ce qui a permis de démontrer que l'attention facilite la perception en modulant les traitements visuels précoces. Quelles en sont les bases neurophysiologiques ? Des études ont démontré que l'activité d'un neurone évoquée par la présentation d'un stimulus dans son champ récepteur augmente lorsque ce stimulus se situe dans le focus attentionnel (e.g. Motter, 1993 ; Reynolds, Pasternak, & Desimone, 2000). Un des mécanismes de facilitation de la perception visuelle par l'attention serait donc d'augmenter la réponse visuelle des neurones.

Dans le cortex visuel, la taille des champs récepteurs augmente avec l'excentricité, leur nombre diminue et l'espacement entre eux augmente. La conséquence d'une telle organisation est que notre résolution spatiale, i.e. notre capacité à discriminer deux points proches dans l'espace, est élevée à la fovéa (le centre de la rétine), et diminue en périphérie. Des expériences comportementales ont démontré que l'attention est capable d'améliorer notre résolution spatiale (pour revue : Anton-Erxleben & Carrasco, 2013 ; Carrasco & Barbot 2014). Nous avons discuté

plus haut d'études en électrophysiologie démontrant que la réponse d'un neurone était biaisée en faveur du stimulus attendu lorsqu'il était présenté dans son champ récepteur (Moran & Desimone, 1985), alors que l'effet du stimulus inattendu sur la réponse du neurone était diminué. Les auteurs ont proposé que cela résulterait d'un rétrécissement et d'un décalage du champ récepteur du neurone vers le stimulus attendu. Cette interprétation a par la suite été confirmée empiriquement grâce à des enregistrements chez le primate non-humain (macaque) de neurones de l'aire MT, sélective au mouvement (Anton-Erxleben, Stephan, & Treue, 2009 ; Womelsdorf, Anton-Erxleben, Pieper, & Treue, 2006). L'attention serait donc capable de refaçonner le profil des champs récepteurs, déplaçant et rétrécissant indépendamment leur centre et leur contour vers la position attendue. Ce mécanisme attentionnel affectant les propriétés des champs récepteurs a aussi été observé pour des stimuli naturels comme des visages (Kay et al., 2015). Un autre mécanisme de facilitation attentionnelle serait donc l'amélioration de la résolution spatiale de notre perception visuelle en modulant de façon dynamique l'organisation de la sensibilité des neurones (pour revue : Anton-Erxleben & Carrasco, 2013). Crucialement, il a été suggéré que si l'attention est capable d'altérer les mécanismes de traitement de traits simples, autrement dit au niveau des aires précoces dans la hiérarchie du système visuel, alors il est probable que l'attention agisse via des connexions descendantes (en anglais *feedback*) depuis des aires de plus haut niveau dans la hiérarchie de traitement de l'information permettant le contrôle attentionnel (e.g. Buschman & Miller, 2009 ; Chelazzi, Miller, Duncan, & Desimone, 1993 ; Dugué, Beck, Marque, & VanRullen, 2019 ; Dugué & VanRullen, 2017 ; Kastner & Pinsk, 2004 ; Gregoriou, Gotts, Zhou, & Desimone, 2009 ; Ibos, Duhamel & Ben Hamed, 2013 ; Miconi & VanRullen, 2016 ; Reynolds & Heeger, 2009 ; Treue, 2001 ; Tsotsos et al., 1995).

CONCLUSION

Comme nous l'avons vu au long de cet article, le système visuel a été largement étudié et décrit du point de vue neurophysiologique. Toutes ces connaissances ont permis de mieux comprendre des règles d'organisation universelles, e.g. les champs récepteurs et leur sélectivité. Ici, nous avons voulu souligner comment cela a aussi permis de mieux comprendre des fonctions cognitives telle que l'attention. L'étude de l'attention, à l'origine du traitement efficace de l'information sensorielle grâce aux phénomènes de sélection et facilitation, a connu des avancées majeures grâce à l'utilisation de la vision comme modèle d'étude. Les connaissances sur le système visuel ont permis de bâtir des protocoles expérimentaux inspirés par sa structure et de mieux comprendre par quels mécanismes l'attention permet de sélectionner l'information et de faciliter le traitement perceptif. C'est en comprenant comment l'attention affecte le système visuel, e.g. en biaisant la compétition entre populations de neurones ou en augmentant l'activité basale d'une population, que nous pouvons mieux comprendre les mécanismes régissant l'attention, ses bases neuronales et ses dysfonctionnements. En dehors de l'attention, la vision a été utilisée pour étudier d'autres aspects de la cognition, comme l'apprentissage par exemple.

De nombreuses caractéristiques de la vision ont été observées de façon universelle et invariante chez des primates humains et non-humains, mais le fonctionnement et la structure de cette modalité sensorielle se développe tout au long de la vie (Grill-Spector, Golarai, & Gabrieli, 2008). Ainsi, l'apprentissage de la lecture réorganise la topographie des aires cérébrales visuelles sélectives aux catégories d'objets (Dehaene et al., 2015). Apprendre à discriminer de nouvelles catégories d'objets réorganise les schémas d'activations dans les aires visuelles occipito-temporales (Op de Beeck, Baker, DiCarlo, & Kanwisher, 2006). Il est donc possible de réorganiser les aires visuelles et leur sélectivité tout au long du développement mais aussi à l'âge adulte, même

pour des catégories d'objets déjà établies (Senoussi, Berry, VanRullen, & Reddy, 2016). L'étendue de la réorganisation du système visuel par de nouveaux apprentissages est corrélée aux performances perceptives dans des tâches effectuées lors de ces apprentissages (McGugin, Gatenby, Gore, & Gauthier, 2012 ; Op de Beeck et al., 2006), mais se généralise aussi à d'autres tâches (Senoussi et al., 2016), attestant ainsi de l'importance centrale de ces changements dans les traitements cognitifs. Ces découvertes sur la manière dont les apprentissages affectent notre comportement et l'organisation cérébrale ont été permises grâce à nos connaissances sur la structure du cortex visuel et son organisation en régions sélectives à des caractéristiques visuelles et des catégories d'objets.

La brève introduction à la vision présentée dans cet article ne fait évidemment pas état de toutes nos connaissances du système visuel et de leur impact sur l'étude des fonctions cognitives. Nous pensons notamment à une autre propriété des neurones de V1 qu'est la binocularité, i.e. la plupart des neurones de V1 combinent les informations provenant de chaque œil (Hubel & Wiesel, 1962). Cette propriété qui permet entre autres de percevoir la profondeur, donne aussi lieu à un phénomène appelé la *rivalité binoculaire* : lorsque l'on présente à un participant deux images différentes à chaque œil, la perception consciente va alterner entre une image et l'autre (Brascamp, Klink, & Levelt, 2015). Ainsi, grâce aux connaissances établies sur l'intégration des projections provenant de chaque œil par les neurones corticaux, et le fait que le système visuel soit organisé de façon rétinotopique, la rivalité binoculaire a permis d'en apprendre davantage sur la perception non-consciente (Engel, Fries, König, Brecht, & Singer, 1999 ; Lee, Blake, & Heeger, 2007). Nous pensons également à une autre propriété du cortex visuel avec son organisation cytoarchitectonique en six couches corticales. Différentes couches ont différentes propriétés physiologiques. Par exemple, certaines reçoivent les projections du CGL, d'autres envoient des projections vers d'autres aires visuelles (e.g. V2, V3) et d'autres encore reçoivent des projections *feedback*. Certains auteurs ont ainsi observé que des informations visuelles sont conservées en mémoire de travail au sein de certaines de ces couches, ou encore que les inférences faites en fonction du contexte, i.e. prédictions sur des stimuli à venir, modulent l'activité des couches superficielles (Gilbert & Sigman, 2007 ; Lawrence et al., 2018 ; Muckli et al., 2015). A une autre échelle, le cortex visuel est intégré au sein d'un réseau fonctionnel local, connexions fonctionnelles entre aires visuelles, et global, connexions avec d'autres aires comme, par exemple, dans le cas du réseau de l'attention (Corbetta & Shulman, 2002 ; Dugué, Merriam, Heeger & Carrasco, 2017, 2018), une telle complexité étant le ciment de la cognition chez l'être humain. En résumé, de nombreuses propriétés connues de la vision sont exploitables pour l'étude des fonctions cognitives.

Le système visuel a néanmoins des particularités, telle que la dynamique temporelle relativement lente des informations qu'il traite, qui peuvent parfois contraindre son utilisation comme modèle d'étude. Contrairement à un flux auditif de parole pour lequel l'identification de la succession rapide de sons est primordiale pour en extraire le sens, une scène visuelle reste relativement stable au cours du temps. Par exemple, la présence et l'agencement d'éléments d'une scène visuelle tels que des arbres, des murs ou le sol, ne varient pas d'une seconde à l'autre. En revanche, dans le cas de l'audition, l'alternance rapide de phonèmes requière une résolution de quelques dizaines de millisecondes pour comprendre le sens des mots. Il a été proposé que la perception opère de façon périodique, i.e. l'information ne serait pas traitée de manière continue, mais comme un flux discret, à la manière d'une caméra vidéo (VanRullen, 2016 ; VanRullen & Koch, 2003). Ainsi, la stabilité de l'information visuelle permettrait une certaine tolérance à son échantillonnage (VanRullen, Zoefel, & Barkin, 2014), i.e. la qualité de notre perception de

l'information visuelle ne serait que peu dégradée par une telle discrétisation (~10Hz). En revanche, imaginons discrétiser un flux de parole à 10Hz, et notre compréhension des phrases s'en verrait significativement affectée (VanRullen, Zoefel, & Barkin, 2014). De récentes études ont toutefois pu mettre en évidence un traitement discret de l'information auditive dans certains cas et suggèrent que ce phénomène n'est pas régi par les mêmes contraintes que dans le système visuel. Il a par exemple été montré que ce sont uniquement les caractéristiques de haut niveau (i.e. syllabes, mots) qui sont traitées périodiquement dans le flux de parole (Zoefel & VanRullen, 2016). L'étude de la nature périodique des processus de traitement sensoriel a donc montré que différentes modalités sensorielles imposent différentes contraintes sur le traitement de l'information, et révèle des mécanismes de traitement différents. Pour conclure, nous argumentons dans cet article que la vision a été et demeure un modèle d'étude particulièrement riche de l'étude de la cognition. Nous soulignons toutefois la nécessité de reproduire les résultats obtenus dans une modalité sensorielle à d'autres modalités, afin de comprendre les mécanismes universels régissant les fonctions cognitives.

Comme nous l'avons déjà mentionné, cet article d'opinion n'a pas la prétention de synthétiser de manière exhaustive le domaine de la vision, ni celui de l'attention, ni même de présenter l'utilisation d'autres modalités sensorielles (e.g. audition) comme modèle d'étude dans les sciences cognitives. Il n'a pas non plus pour objectif d'aborder tous les domaines des sciences cognitives – ici, nous nous sommes principalement concentrés sur les neurosciences et la psychologie cognitive, mais avons aussi brièvement abordé la neuropsychologie, la modélisation et la philosophie. Grâce à une série d'exemples, nous avons argumenté que les systèmes sensoriels peuvent être utilisés comme modèles d'étude de la cognition dans divers domaines des sciences cognitives. En particulier, l'accumulation de nos connaissances de l'organisation structurelle caractéristique du système visuel – notamment la rétinotopie, les champs récepteurs et la notion de sélectivité – en font un modèle particulièrement riche de l'étude des fonctions cognitives telles que l'attention. L'exploitation de ces connaissances sur l'organisation physiologique a permis de mieux décrire et appréhender les mécanismes psychologiques sous-jacents à la cognition humaine.

RÉFÉRENCES

- Alonso, J.-M., & Chen, Y. (2009). Receptive field. *Scholarpedia*, 4(1), 5393.
- Anton-Erxleben, K., & Carrasco, M. (2013). Attentional enhancement of spatial resolution: Linking behavioural and neurophysiological evidence. *Nature Reviews Neuroscience*, 14(3), 188–200.
- Anton-Erxleben, K., Stephan, V. M., & Treue, S. (2009). Attention Reshapes Center-Surround Receptive Field Structure in Macaque Cortical Area MT. *Cerebral Cortex*, 19(10), 2466–2478.
- Baizer, J. S., Ungerleider, L. G., & Desimone, R. (1991). Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience*, 11(1), 168–190.
- Bar-Cohen, Y. (2006). Biomimetics—using nature to inspire human innovation. *Bioinspiration & biomimetics*, 1(1), P1
- Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The 'when' pathway of the right parietal lobe. *Trends in cognitive sciences*, 11(5), 204-210.
- Bergmann, J., Genç, E., Kohler, A., Singer, W., & Pearson, J. (2016). Smaller Primary Visual Cortex Is Associated with Stronger, but Less Precise Mental Imagery. *Cerebral Cortex*, 26(9), 3838–3850.

- Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Springer
- Block, N. (2010). Attention and Mental Paint. *Philosophical Issues*, 20(1), 23–63.
- Brascamp, J. W., Klink, P. C., & Levelt, W. J. M. (2015). The ‘laws’ of binocular rivalry: 50 years of Levelt’s propositions. *Vision Research*, 109, 20–37.
- Breasted, J. H. (1930). *The Edwin Smith Surgical Papyrus: Published in facsimile and hieroglyphic transliteration with translation and commentary in two volumes* (Vol. 3). University of Chicago Press Chicago.
- Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the “spotlight” of visual attention. *Nature Neuroscience*, 2(4), 370–374.
- Buschman, T. J., & Miller, E. K. (2009). Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron*, 63(3), 386–396.
- Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1), 51–62.
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51(13), 1484–1525.
- Carrasco, M., & Barbot, A. (2014). How attention affects spatial resolution. In Cold Spring Harbor Symposia on Quantitative Biology (Vol. 79, pp. 149–160). Cold Spring Harbor Laboratory Press.
- Carrasco, M., & Barbot, A. (2019). Spatial attention alters visual appearance. *Current opinion in psychology*, 29(56–64), 1.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363(6427), 345–347.
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area. *Brain*, 123(2), 291–307.
- Collins, T., Andler, D., & Tallon-Baudry, C. (2018). *La cognition : Du neurone à la société*. Gallimard.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3), 201.
- Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, 16(4), 234–244.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A Cortical Area Selective for Visual Processing of the Human Body. *Science*, 293(5539), 2470–2473.
- Dugué, L., Beck, A. A., Marque, P., & VanRullen, R. (2019). Contribution of FEF to attentional periodicity during visual search: A TMS study. *ENeuro*.
- Dugué, L., Marque, P., & VanRullen, R. (2011). Transcranial magnetic stimulation reveals attentional feedback to area V1 during serial visual search. *PloS One*, 6(5), e19712.
- Dugué, L., Marque, P., & VanRullen, R. (2015). Theta Oscillations Modulate Attentional Search Performance Periodically. *Journal of Cognitive Neuroscience*, 27(5), 945–958.
- Dugué, L., Merriam, E. P., Heeger, D. J., & Carrasco, M. (2017). Specific visual sub-regions of TPJ mediate reorienting of spatial attention. *Cerebral Cortex*.
- Dugué, L., Roberts, M., & Carrasco, M. (2016). Attention Reorients Periodically. *Current Biology*, 26(12), 1595–1601.
- Dugué, L., & VanRullen, R. (2017). Transcranial Magnetic Stimulation Reveals Intrinsic Perceptual and Attentional Rhythms. *Frontiers in Neuroscience*, 11, 154.
- Dugue, L., Merriam, E. P., Heeger, D. J., & Carrasco, M. (2018). Endogenous and exogenous

- attention distinctly modulate fMRI activity in visual cortex. *bioRxiv*, 414508.
- Eichenbaum, H. (1998). Using Olfaction to Study Memory. *Annals of the New York Academy of Sciences*, 855(1), 657–669.
- Eickenberg, M., Gramfort, A., Varoquaux, G., & Thirion, B. (2017). Seeing it all: Convolutional network layers map the function of the human visual system. *NeuroImage*, 152, 184–194.
- Eickenberg, M., Senoussi, M., Pedregosa, F., Gramfort, A., & Thirion, B. (2013). Second Order Scattering Descriptors Predict fMRI Activity Due to Visual Textures. *2013 International Workshop on Pattern Recognition in Neuroimaging*, 5–8.
- Engel, A. K., Fries, P., König, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Consciousness and Cognition*, 8(2), 128–151.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598–601.
- Feinberg, T. E., & Farah, M. J. (2000). *Patient-based Approaches to Cognitive Neuroscience*. MIT Press.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed Hierarchical Processing in the Primate. *Cerebral Cortex*, 1(1), 1–47.
- Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, 14(9), 1195–1201.
- Gilbert, C. D., & Sigman, M. (2007). Brain States: Top-Down Influences in Sensory Processing. *Neuron*, 54(5), 677–696.
- Goodale, M. A., Meenan, J. P., Bühlhoff, H. H., Nicolle, D. A., Murphy, K. J., & Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology*, 4(7), 604–610.
- Grill-Spector, K., Golarai, G., & Gabrieli, J. (2008). Developmental neuroimaging of the human ventral visual cortex. *Trends in Cognitive Sciences*, 12(4), 152–162.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *Journal of Neurophysiology*, 35(1), 96–111.
- Gross, C. G. (1999). *Brain, Vision, Memory: Tales in the History of Neuroscience*. MIT Press.
- Güçlü, U., & Gerven, M. A. J. van. (2015). Deep Neural Networks Reveal a Gradient in the Complexity of Neural Representations across the Ventral Stream. *Journal of Neuroscience*, 35(27), 10005–10014.
- Handel, S. (1989). Listening: An introduction to the perception of auditory events.
- Hartline, H. K. (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *American Journal of Physiology-Legacy Content*, 121(2), 400–415.
- He, S., & Davis, W. L. (2001). Filling-in at the natural blind spot contributes to binocular rivalry. *Vision research*, 41(7), 835–840.
- Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (1996). Computational models of cortical visual processing. *Proceedings of the National Academy of Sciences*, 93(2), 623–627.
- Henschen, S. E. (1893). On the visual path and centre. *Brain*, 16(1–2), 170–180.
- Herrmann, K., Montaser-Kouhsari, L., Carrasco, M., & Heeger, D. J. (2010). When size matters: Attention affects performance by contrast or response gain. *Nature Neuroscience*, 13(12), 1554–1559.
- Holmes, G., & Lister, W. T. (1916). Disturbances of vision from cerebral lesions, with special reference to the cortical representation of the macula. *Brain*, 39(1–2), 34–73.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex.

- The Journal of Physiology*, 148(3), 574–591.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, 160, 106–154.
- Ibos, G., Duhamel, J. R., & Hamed, S. B. (2013). A functional hierarchy within the parietofrontal network in stimulus selection and attention control. *Journal of Neuroscience*, 33(19), 8359–8369.
- Inouye, T. (1909). Die Sehstörungen bei Schussverletzungen der kortikalen Sehsphäre. *Nach Beobachtungen an Verwundeten Der Letzten Japanischen Kriege*.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience*, 17(11), 4302–4311.
- Kastner, S., & Pinsk, M. A. (2004). Visual attention as a multilevel selection process. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 483–500.
- Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature*, 452(7185), 352–355.
- Kay, K. N., Weiner, K. S., & Grill-Spector, K. (2015). Attention Reduces Spatial Uncertainty in Human Ventral Temporal Cortex. *Current Biology*, 25(5), 595–600.
- Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. *PLOS Computational Biology*, 10(11), e1003915.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25 (Lake Tahoe, NV)*, 1097–1105.
- Lawrence, S. J. D., van Mourik, T., Kok, P., Koopmans, P. J., Norris, D. G., & de Lange, F. P. (2018). Laminar Organization of Working Memory Signals in Human Visual Cortex. *Current Biology*, 28(21), 3435–3440.e4.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- Lee, S.-H., Blake, R., & Heeger, D. J. (2007). Hierarchy of cortical responses underlying binocular rivalry. *Nature Neuroscience*, 10(8), 1048–1054.
- Lennie, P. (2003). The Cost of Cortical Computation. *Current Biology*, 13(6), 493–497.
- Liu, T., Pestilli, F., & Carrasco, M. (2005). Transient attention enhances perceptual performance and fMRI response in human visual cortex. *Neuron*, 45(3), 469–477.
- Long, B., Yu, C.-P., & Konkle, T. (2018). Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proceedings of the National Academy of Sciences*, 115(38), E9015–E9024.
- Marčelja, S. (1980). Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.*, 70(11), 1297–1300.
- Marr, D. (2010). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press.
- McGugin, R. W., Gatenby, J. C., Gore, J. C., & Gauthier, I. (2012). High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance. *Proceedings of the National Academy of Sciences*, 109(42), 17063–17068.
- Miconi, T., & VanRullen, R. (2016). A Feedback Model of Attention Explains the Diverse Effects of Attention on Neural Firing Rates and Receptive Field Structure. *PLOS Computational*

- Biology*, 12(2), e1004770.
- Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, 6(1), 57–77.
- Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, 6(0), 414–417.
- Monier, C. (2018). Neuroscience in the Sciences of Cognition: A Review of the Literature Arguing for Neuroscience to Becoming Spinozist. In Monier Cyril & Sarti Alessandro (Eds), *Neuroscience In The Sciences of Cognition - between Neuroenthusiasm and Neuroskepticism*, *Intellectica*, 69, (pp.27-132)
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229(4715), 782–784.
- Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, 70(3), 909–919.
- Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., ... Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Current Biology*, 25(20), 2690–2695.
- Nachmias, J. (1967). Effect of Exposure Duration on Visual Contrast Sensitivity with Square-Wave Gratings. *JOSA*, 57(3), 421–427.
- Nobre, A. C., & Kastner, S. (Eds.) (2014). *The Oxford handbook of attention*. Oxford University Press.
- Nowak, L. G., & Bullier, J. (1997). The Timing of Information Transfer in the Visual System. In K. S. Rockland, J. H. Kaas, & A. Peters (Eds.), *Extrastriate Cortex in Primates* (pp. 205–241).
- Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination Training Alters Object Representations in Human Extrastriate Cortex. *The Journal of Neuroscience*, 26(50), 13025–13036.
- Park, W. J., Schauder, K. B., Zhang, R., Bennetto, L., & Tadin, D. (2017). High internal noise and poor external noise filtering characterize perception in autism spectrum disorder. *Scientific reports*, 7(1), 17584.
- Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nat Neurosci*, 5(12), 1332–1338.
- Pearson, J. (2019). The human imagination: the cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, 1-11.
- Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in Neurobiology*, 18(4), 445–451.
- Pestilli, F., Carrasco, M., Heeger, D. J., & Gardner, J. L. (2011). Attentional enhancement via selection and pooling of early sensory responses in human visual cortex. *Neuron*, 72(5), 832–846.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, 109(2), 160–174.
- Rajimehr, R., Bilenko, N. Y., Vanduffel, W., & Tootell, R. B. H. (2014). Retinotopy versus Face Selectivity in Macaque Visual Cortex. *Journal of Cognitive Neuroscience*, 26(12), 2691–2700.
- Reynolds, J. H., & Heeger, D. J. (2009). The Normalization Model of Attention. *Neuron*, 61(2),

- 168–185.
- Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention Increases Sensitivity of V4 Neurons. *Neuron*, 26(3), 703–714.
- Robertson, C. E., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience*, 18(11), 671–684.
- Rosenblatt, F. (1962). *Principles of neurodynamics; perceptrons and the theory of brain mechanisms*. No. VG-1196-G-8. Cornell Aeronautical Lab Inc Buffalo NY, 1961.
- Rossetti, Y., Pisella, L., & Vighetto, A. (2003). Optic ataxia revisited. *Experimental Brain Research*, 153(2), 171–179.
- Schneider, G. E. (1969). Two visual systems. *Science (New York, N.Y.)*, 163(3870), 895–902.
- Senoussi, M., Berry, I., VanRullen, R., & Reddy, L. (2016). Multivoxel Object Representations in Adult Human Visual Cortex Are Flexible: An Associative Learning Study. *Journal of Cognitive Neuroscience*, 28, 852–868.
- Senoussi, M., Moreland, J. C., Busch, N. A., & Dugué, L. (2019). Attention explores space periodically at the theta frequency. *Journal of Vision*, 19(5), 22–22.
- Sereno, M. I., McDonald, C. T., & Allman, J. M. (1994). Analysis of Retinotopic Maps in Extrastriate Cortex. *Cerebral Cortex*, 4(6), 601–620.
- Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, 104(15), 6424–6429.
- Sherrington, S. C. S. (1906). *The Integrative action of the nervous system*. Yale University Press.
- Skottun, B. C., Bradley, A., Sclar, G., Ohzawa, I., & Freeman, R. D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: A comparison of single cells and behavior. *Journal of Neurophysiology*, 57(3), 773–786.
- Stevens, K. A. (2012). The vision of David Marr. *Perception*, 41(9), 1061–1072.
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015 (pp. 1-9).
- Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J.-B., Lebihan, D., & Dehaene, S. (2006). Inverse retinotopy: Inferring the visual content of images from brain activation patterns. *NeuroImage*, 33(4), 1104–1116.
- Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences*, 24(5), 295–300.
- Tsotsos, J. K., Culhane, S. M., Kei Wai, W. Y., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, 78(1), 507–545.
- Van Essen, D. C. (2003). Organization of visual areas in macaque and human cerebral cortex. *The Visual Neurosciences*, 1, 507–521.
- VanRullen, R., Zoefel, B., & Barkin, I. (2014). On the cyclic nature of perception in vision versus audition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1641), 20130214.
- VanRullen, R., Delorme, A., & Thorpe, S. (2001). Feed-forward contour integration in primary visual cortex based on asynchronous spike propagation. *Neurocomputing*, 38–40, 1003–1009.
- Wandell, B. A., & Winawer, J. (2011). Imaging retinotopic maps in the human brain. *Vision Research*, 51(7), 718–737.
- Womelsdorf, T., Anton-Erxleben, K., Pieper, F., & Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nature Neuroscience*, 9(9), 1156–

1160.

- Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences*, *111*(23), 8619–8624.
- Zoefel, B., & VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage*, *124*, 16–23.